Physica A 391 (2012) 5244–5253

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

# Users structure and behavior on an online social network during a political protest

A.J. Morales, J.C. Losada, R.M. Benito *

*Grupo de Sistemas Complejos, Universidad Politécnica de Madrid, ETSI Agrónomos, 28040, Madrid, Spain*

ARTICLE INFO

*Article history:*
Received 19 December 2011
Received in revised form 29 February 2012
Available online 31 May 2012

*Keywords:*
Complex networks
Social behavior
Information flow
Twitter

ABSTRACT

Over the past years, new technologies and specially online social networks have penetrated into the world's population at an accelerated pace. In this paper we analyze collected data from the web application Twitter, in order to describe the structure and dynamics of the emergent social networks, based on complexity science. We focused on a Venezuelan protest that took place exclusively by Twitter during December, 2010. We found a community structure with highly connected hubs and three different kinds of user behavior that determine the information flow dynamics. We noticed that even though online social networks appear to be a pure social environment, traditional media still holds loads of influence inside the network.



## 1. Introduction

Online social networks have gained an enormous influence among the global population over the last years. There are several types of these web applications, each with its own purpose, such as Facebook for personal relations, Youtube for sharing videos or Twitter to exchange text messages. Such is the speed and facility that provides these communication tools, that it is indeed revolutionizing marketing techniques [1], cultural events, political activism [2] and emergencies management. An important fact of these technologies is that they provide an enormous amount of user generated data which may be collected and analyzed, bringing lots of opportunities for human behavior research [3].

In this context Twitter introduces itself as a fast information diffusion network, that allows people to be informed about ongoing events in real time and to interact with other users. There are several usage motivations for this network [4], but lately it has been used as a platform for social collective emergence and coordination, like the Arab Spring, where it played an important role in the event's development. As a result, Twitter is allowing people to become an active part in ongoing stories, leaving behind the passivity associated to radio, television or press. For researchers it is an opportunity to analyze the social structures that emerge from the users' relations [5,6] and, because actions are usually time stamped, it also allows to study social dynamics, like influence dynamics [7–10], communities interactions [11] and prediction of information cascades [12].

Our main goal is to describe and analyze the effects that the user behavior have over the social structure emergence and the information flow dynamics on the Twitter social network. For this matter we have carried out a quantitative analysis of a Venezuelan virtual protest that took place during December 2010, exclusively on Twitter. We built two networks to represent the phenomena. On one hand, we constructed the social substratum in which the information may have flowed, and on the other hand, the network where the messages were retransmitted and the information actually flowed. Then, based on the graph theory [13], we have calculated and correlated several measures to understand the social structure and the dynamical patterns that emerge from the studied conversation. In summary, we found that people organize in a community structure around highly connected hubs which play lots of influence inside the network, mainly formed by three different kinds of user with highly asymmetrical profiles that determine the information flow dynamics.

* Corresponding author. Tel.: +34 913365646; fax: +34 913365726.
*E-mail addresses:* alfredo.moralesg@alumnos.upm.es (A.J. Morales), rosamaria.benito@upm.es (R.M. Benito).

0378-4371/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2012.05.015

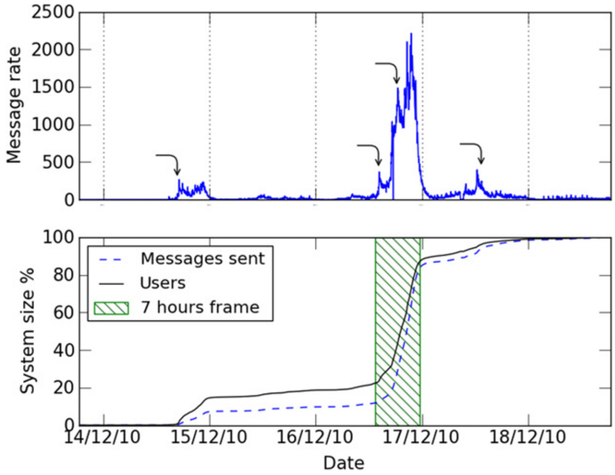


**Fig. 1.** Top: Temporal variation of the message rate (messages/min) of the Twitter Venezuelan protest #SOSInternetVE. Arrows indicate some of the times when the protest convoker participated. Bottom: System growth through time in terms of the percentage of messages (dashed line) and the percentage of participant users (solid line).

The organization of this paper is as follows. In Section 2 we describe the system under study. We begin by describing the web application Twitter, from where we have collected the data, and then we describe the main statistical characteristics of the protest. Then, in Section 3, we describe the structures formed by the users when they interact with each other, either passively or actively. Next, in Section 4, we describe the user behavior and dynamics that produce such structures. And finally in Section 5 we describe these structures from the mesoscale point of view.

## 2. System

This work is based on the online social network Twitter. This web application has over 200 million users all around the globe and keeps growing everyday. Its main feature consists in allowing people to post and exchange text messages limited to 140 characters. It also allows people to follow other users whose messages are broadcast among its followers in real time. The Twitter's global followers network is a directed graph where non reciprocal relations are admitted. An important mechanism on Twitter is the retweet, where users retransmit read messages to their own followers. This mechanism allows individual messages to propagate and travel throughout the network. Also, messages may be identified using keywords called hashtags. This mechanism generates the trending topics, and people use it to discuss and exchange ideas without the necessity of having any explicit relation between them.

Our dataset is constructed from public access messages posted on Twitter, related to a Venezuelan protest that took place exclusively by digital means. The event consisted in posting messages identified with the hashtag #SOSInternetVE on December 16, 2010. We downloaded all the messages that included the hashtag #SOSInternetVE using the Twitter API interface between December 14–19, 2010. Each message contains information about its author, creation date, device source and text body. In total we found 421,602 messages, written by 77,706 users. It is remarkable that 42% of messages where retweets and 60% were sent from smart mobile phones. This last result is in accordance with the latest user tendencies research [14], where is stated that most of users participate in the online social media away from personal computers.

At the top of Fig. 1 we present the evolution of the message rate between December 14–19, 2010, which has a similar shape as the Twitter time series modeled in the study of Yang [15]. It can be noticed that in the beginning of December 14, 2010, the studied hashtag did not even exist on the Twitter servers. Then, after its first appearance on the same day, some user activity was recorded. Yet it is on December 16, 2010, when the protest takes actual place and the trending topic bursts and reaches its highest point, showing critical phenomena features. However, after December 18, 2010, much of the interest is lost and the trend tends to decay really fast as expected for trend topics on Twitter [16].

The protest growth can be seen more clearly at the bottom of Fig. 1, where we have plotted the accumulated number of messages (dashed line) and users (solid line) as a function of time. It is remarkable that the system grew from 22% to 87%, in terms of users, and 12% to 84%, in terms of messages, in a time frame of 7 h, which has been highlighted around the afternoon of December 16, 2010 in Fig. 1, and coincides with the main burst. Furthermore, it can be noticed that the number of users that participate in the protest saturates much faster than the amount of messages at all times. This is a typical feature of local interest conversations [5] where users post messages repetitively on the same topic. For example, after the day the protest was convoked on December 14, 2010, already 15% of the users had participated. However, the messages they posted did not even reach 7% of the total amount.

In Fig. 2 we show the evolution of the cumulative distribution of the number of messages sent (posted) by users, on the different days that the protest lasted. It can be noticed that the distribution can be fitted to an exponentially truncated power law, in the form: $P(x > x^*) \propto x^{-\beta} e^{-x/c}$, where $\beta = 0.88 \pm 1.2 \times 10^{-3}$ and $c = 65 \pm 0.59$ on the last day. It is remarkable

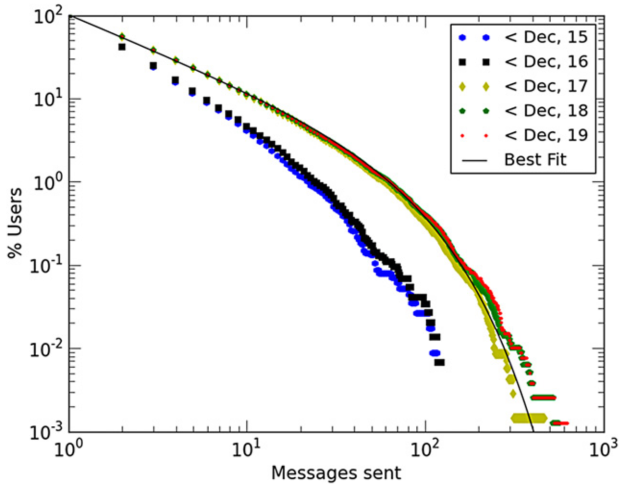


**Fig. 2.** User cumulative distribution according to the number of sent messages through time. The solid line is the fit to an exponentially truncated power law, $P(x > x^*) \propto x^{-\beta} e^{-x/c}$, where $\beta = 0.88 \pm 1.2 \times 10^{-3}$ and $c = 65 \pm 0.59$ on the last day.

that there is a clear distinction between the days before and after the main burst (see Fig. 1) which reflects the criticality of the phenomena. However, on each day of both stages, the users presented the same behavior, in the sense that they are distributed in the same way during the days before the protest, but also during the days after the protest.

This distribution indicates a certain degree of complexity in the phenomena and heterogeneity in the user behavior. Before the main burst, 60% of the participants had sent less than a couple of messages, 1% over 30 messages, and about 0.01% had posted over 100 messages. On the other hand, on the last day of the protest, 50% of the users also had sent a couple of messages at most, while 1% sent over 60 messages, and just about 0.0013% posted over 600 messages. This result shows that the percentage of most active users decreases rapidly as the system grows.

## 3. Networks

The purpose of this study is to analyze the users' roles in the information diffusion dynamic during the protest. For this matter we built two networks based on the previous mentioned Twitter relations. First, we constructed a network to represent the fraction of the global Twitter follower graph that participated in the protest, that we have identified as the social substratum at which the information may flow. Second, we built the information diffusion network relating the users who retransmitted messages and were retransmitted, to represent the effective channels through which the messages actually traveled inside the social substratum. In both cases, edges are not destroyed as time goes on, so the data delivered represents the accumulated interactions at the end of the protest.

The first one is a network based on who follows who. This relation indicates who received whose messages. As said before, it represents a subgraph of the global Twitter social network, built exclusively with the participant users in this specific conversation. As nonreciprocal relations are admitted in the application, the resulting is an asymmetric graph, where the edges are directed and non weighted. At the end of the protests on December 19, 2010 the network is compound by 77,706 nodes and 5,761,331 edges. The nodes are the users who participated in the protest and the edges are drawn from each followers lists. We have included all the participants in our calculations.

The nodes' in degree indicates how many people follow a certain user, or the equivalent, how many people received the user's messages. On the other hand, the out degree means the number of people that a certain user follows, which indicates from how many people it has received messages. Both in and out degrees follow power law distributions as shown in Fig. 3. In terms of the in degree, the distribution indicates that over 50% of the users are followed by less than 15 users, while just 1% of the users have over 1000 followers and around 0.01% of the users have over 20,000 followers. For the out degree distribution, we found that over 50% of the users follow less than 40 users, while 1% of the participants follow over 600 users and 0.01% follow over 9,500 users. This distribution presents an exponent within the expected range for human actions, as it has been stated by Newman in Ref. [17].

As can be seen in Table 1, the mean distance between nodes in this network is $d_F = 2.2$. This value indicates the presence of the small world effect [18], which was first reported by Stanley Milgram in his famous experiment where he detected that on average two randomly chosen people could be linked through 6 hops (intermediary persons) from each other [19]. Previous studies performed on the Twitter global follower graph, state that the mean distance between users is to be 4.12 [4]. This fact is usually related to the presence of users that act like hubs, concentrating a large quantity of incoming and outgoing links. However, our results are lower than the previously reported values, due to the special characteristics of the event and its participants. For example, the protest convoker, which is a TV station, is followed by over 52% of the participants, linking half of the total population.

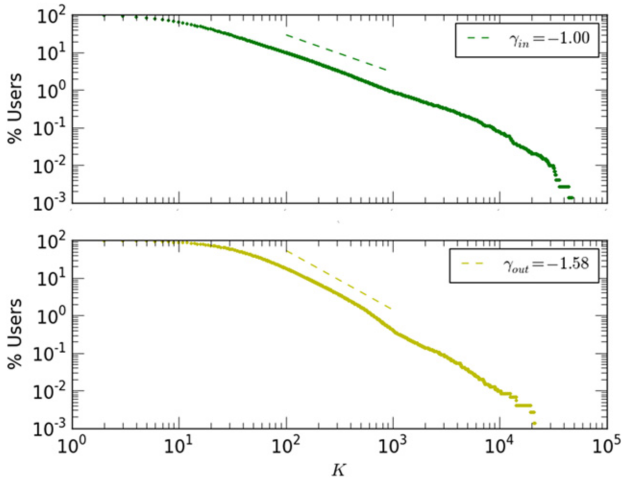


**Fig. 3.** In (top) and out (bottom) degree cumulative distributions of the followers network.

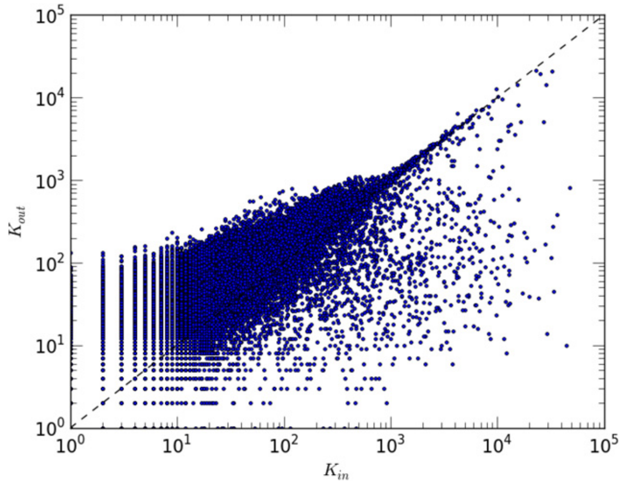


**Fig. 4.** In and out degrees correlation of the followers network.

**Table 1**
Followers and retransmission network parameters.

| Network | Nodes | Edges | Mean distance | Density |
|---|---|---|---|---|
| Followers | 77,706 | 5,761,331 | 2.22 | 1.42E−3 |
| Retransmissions | 54,423 | 231,485 | 3.40 | 1.25E−4 |

Based on the degrees correlation shown in Fig. 4, we found that user profiles are highly heterogeneous and that the network is very asymmetrical. It is remarkable that there are some users, corresponding to the scattered points located below the dotted diagonal, that are widely followed but do not follow many people. At the same time we found other users, who are more reciprocal and stay near the dotted diagonal, specially after $K_{in} > 1000$ followers, where practically any users are found above the diagonal. Finally there are some users, corresponding to the region densely located above the diagonal, who follow more people than those following them. These users represent the majority of the participants.

The second network is built according to who retransmits whose messages. It is a network that emerges from the users' interactions. The nodes are users that retransmitted messages to its own followers, as well as users whose messages were retransmitted. This network indicates the effective links through which the information actually flows inside the active social substratum. In principle it might seem to be a subgraph of the follower networks, but it is not so, since on Twitter people are able to retransmit any message, no matter if it does not hold any type of relation with the source user. The resulting network is also a directed graph, where edges are weighted according to the number of times a user retweeted the source user. In total, by December 19, 2010, the graph is compounded by 54,423 nodes and 231,485 links. The difference between

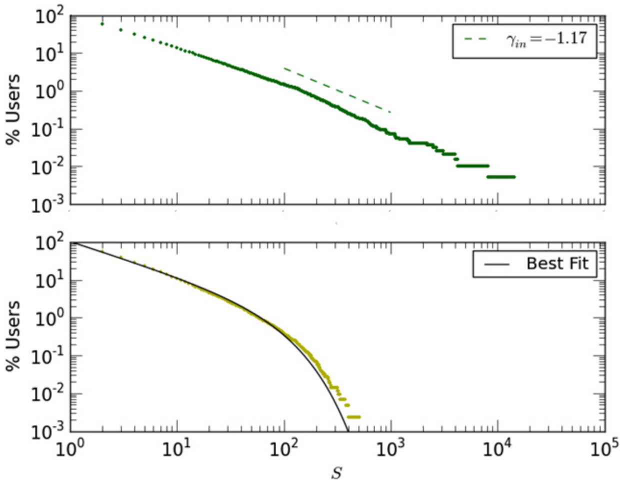


**Fig. 5.** In (top) and out (bottom) strength cumulative distributions of the retransmission network. The solid line is the fit to an exponentially truncated power law $P(S_{out} > S^*_{out}) \propto S_{out}^{-\beta} e^{-S_{out}/c}$, where $\beta = 0.89 \pm 1.7 \times 10^{-3}$ and $c = 61 \pm 1.15$.

the amount of nodes found in the followers graph, shows that 30% of the users behave much more passively than the others. Furthermore, we found that 75% of the participants were not retweeted at all.

In the retransmissions network, we have analyzed the strength function for each user. The in strength value represents the number of times a user has been retweeted. Its distribution follows a power law, as shown at the top of Fig. 5. Such a distribution indicates the presence of highly connected hubs, which explains why the mean distance between nodes is $d_R = 3.4$, which is also a very low value. On the other hand, the out strength shows the number of times a single user has retransmitted. Its distribution can be fitted better to an exponentially truncated power law distribution, as shown at the bottom of Fig. 5. The truncation value, near 500, is related to the limitation for human actions as stated in the Dumbar number theory [20]. This theory states that people are only able to maintain tie relationships with less than 200 people. The reason for which we found a higher value relies on the fact that a retweet does not imply strictly a mutual relation between people. In fact, it is an individual choice that has a very low cost in money, time and personal energy, which makes it easy to happen.

The difference between the in and out strength distributions, is related to the way that we have designed the network. While the out strength is due to one person's activity, the in strength distribution is due to the aggregation of several individual efforts. Such aggregation is responsible for the emergence of extreme cases and a higher complexity level in the final distribution. From the in strength distribution, shown at the top of Fig. 5, it can be noticed that over 60% of the users that participated in the retransmission process gained less that 3 retransmissions, while 1% gained more than 150 retransmissions, and only 0.01% gained over 5000 retransmission. Analogously, for the out strength distribution, we found that over 60% of the users who retransmitted messages, did it over less than 3 messages, while 1% of them retweeted over 60 messages, and less than 0.01% retransmitted more than 300 messages.

We also calculated the edge's weight distribution and found that it follows a power law as shown in Fig. 6. The edge's weight represents the number of times that a single user retweeted another user. The figure shows that only 10% of the edges present a weight higher than 2. However, we found that near 0.001% of the edges have a weight higher than 80. This indicates that the majority of users retweeted other users individually only a couple of times, yet a small fraction of them maintained a closer tie with other users, in the sense that they retweeted their messages close to 100 times. On the other hand, the retransmission network also presents the same asymmetries found in the followers network. For example the 10 most retransmitted users caused more than 20% of all retransmissions, writing less than 0.4% of all messages.

It is remarkable that the retransmissions network is much less dense than the followers network as stated in Table 1. This indicates that inside the contacts web there is a finer structure where the information actually travels. The reason for this result is that retransmitting implies an active behavior, instead of the passivity of the following relation. This shows how users are more selective when it comes to take some action.

## 4. Analysis of user behavior

So far, we have discussed that people's profiles are very heterogeneous according to the audience and source sizes, retransmission levels and amount of messages posted, whereas they are original or retransmitted from others. In Fig. 7 we present the in strength, $S_{in}$, of the retransmissions network as a function of the relation between the in and out degrees, $K_{in}/K_{out}$, of the followers network. The users are represented by points colored by the users' activity or amount of messages posted. It is important to clear out that the users that appear in this plot were retransmitted at least once. These users represent 25% of the participants, as said in Section 3.

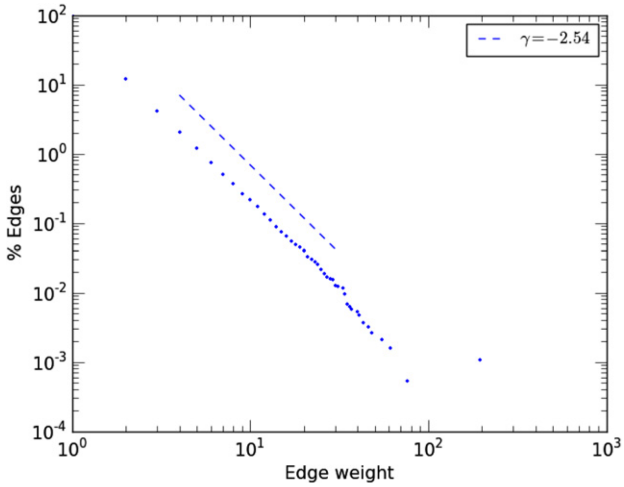


**Fig. 6.** Edge's weight cumulative distribution of the retransmission network.

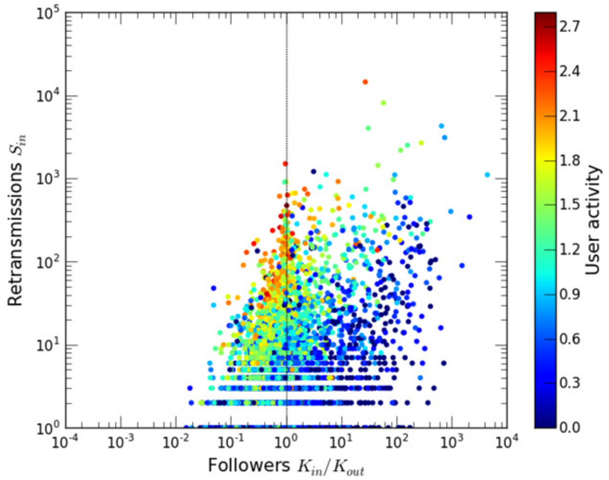


**Fig. 7.** Retransmission in strength versus the ratio of followers in and out degrees. Color is related to the user activity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

This presentation let us separate the popular accounts, where $K_{in}/K_{out} > 1$, from the non popular accounts, where $K_{in}/K_{out} < 1$, and the reciprocal users, where $K_{in}/K_{out} \sim 1$. It can be noticed that the popular accounts may get a high value of retransmissions while having low activity. Meanwhile, the reciprocal users also get the same amount of retransmissions compared to the popular accounts, but they must employ much more activity. Yet if a popular account increases its activity, the retransmission level boosts nonlinearly, like the most retransmitted user that gained more than 10,000 retransmissions.

This result let us classify users into three categories: Information producers, active consumers and passive consumers. The information producers are the widely followed users who gain an enormous amount of retransmissions, whereas they have low activity. These users do not tend to follow a lot of people, nor retransmit many messages. We found that these accounts belong to traditional mass media agents like TV, journalists, politicians and celebrities. On the other hand active consumers are users with high reciprocity in relations. They tend to gain as much audience and retransmission rate, as the amount of activity employed. They are key in the information diffusion process, because they boost the content and serve as the propagators of the information producers. At last, passive consumers are the largest group of users who practically does not participate in the propagation process. They consume more information than what they produce. They are characterized for having low activity rate, not retransmitting many messages and receiving messages from many more people than their audiences.

## 5. Mesoscale communities

In order to get more insight in the structure and behavior of the Twitter users during the protest, we have calculated the mesoscale structure for both networks. It has been shown that complex networks from social and biological phenomena

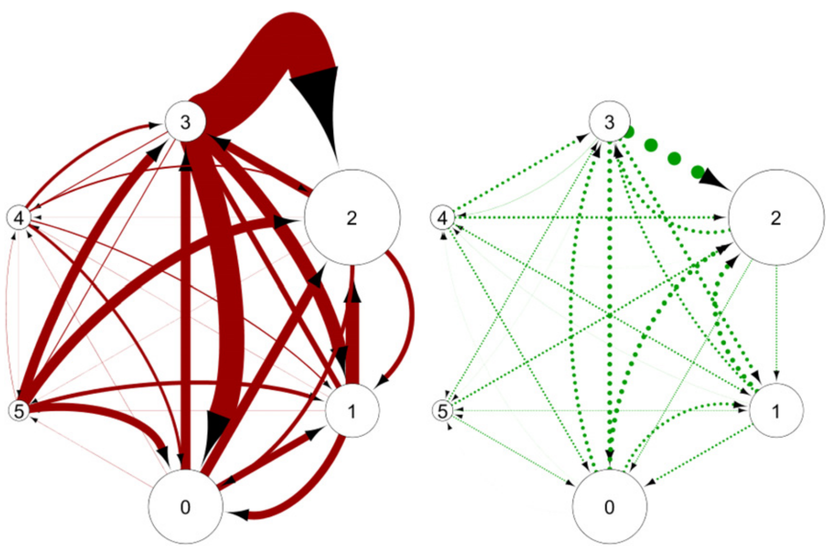


**Fig. 8.** Community structure for the follower graph. Circles represent communities of users and their size is proportional to the amount of users that belong to the community. Edges represent the inter-community links, either followers (Left) or retransmissions (Right), and their width is proportional to the amount of edges, normalized by the size of the outgoing community.

**Table 2**
Main collectives around which each follower community is formed.

| Community | Collective |
|---|---|
| 0 | Comedy accounts |
| 1 | Show business celebrities |
| 2 | Opposition media |
| 3 | Opposition politicians |
| 4 | International media |
| 5 | Government favorable politicians and media |

usually present nodes which are more densely connected between them than with the rest of the network [21]. In this section we describe the communities detected in our networks based on the algorithm described in Ref. [22]. We chose this algorithm based on the modularity optimization, due to its capacity to reveal mesoscale structure in large graphs with good computing performance.

On the follower graph, we found six main communities that grouped over 98% of the population. We identified the most followed users at each community in order to understand the reasons for which people have grouped. We found that these structures are formed by users around central accounts that belong to similar collectives. Specifically, we found communities around opposition media and journalists, opposition politicians, entertainment celebrities, international media, comedy accounts and government favorable politicians, as described in Table 2.

This structure is shown on the left side of Fig. 8. Each node represents a different community and its size is proportional to the amount of users that compounds them. We found that the largest communities are formed around the comedy accounts (0), celebrities (1), opposition media and journalists (2) and opposition politicians (3). Meanwhile, the smallest ones are formed around international media (4) and government favorable users (5). The edges shown in Fig. 8 represent the inter-community links, and indicates the existence of users who follow or are followed by users from other communities. The edge width is proportional to the amount of individual inter-community links, normalized by the size of the outgoing community. As it has been pointed out in Section 2, messages go from the source to its followers. Thus the information flows in the opposite sense of the edges.

It can be noticed that there is a tie relation between the communities formed around the opposition media, opposition politicians, celebrities and comedy accounts. These collectives seem to have dominated the protest. Specially the opposition media community, group 2 in Fig. 8, which concentrates the most amount of users and incoming links. Therefore their messages are widely received throughout the network. Group 3 is certainly smaller than other groups, which is a remarkable fact because it concentrates much of the opposition politicians and the event consisted of an opposition political protest. However, this group is strongly related with other communities and, even though they present a large amount of outgoing links, their messages are also quite spread. In contrast, group 5, which represents the government favorable accounts, seems to follow a lot of outside users, yet only a little fraction of the participants seem to follow them. This means that most of their messages mainly remain inside their community and are hardly read by the rest of the network. Nevertheless, for all communities we found the same user behavior. In the sense that all of them are formed around popular accounts that belong to traditional mass media agents, no matter if they are opposition or government favorable, or even non Venezuelan

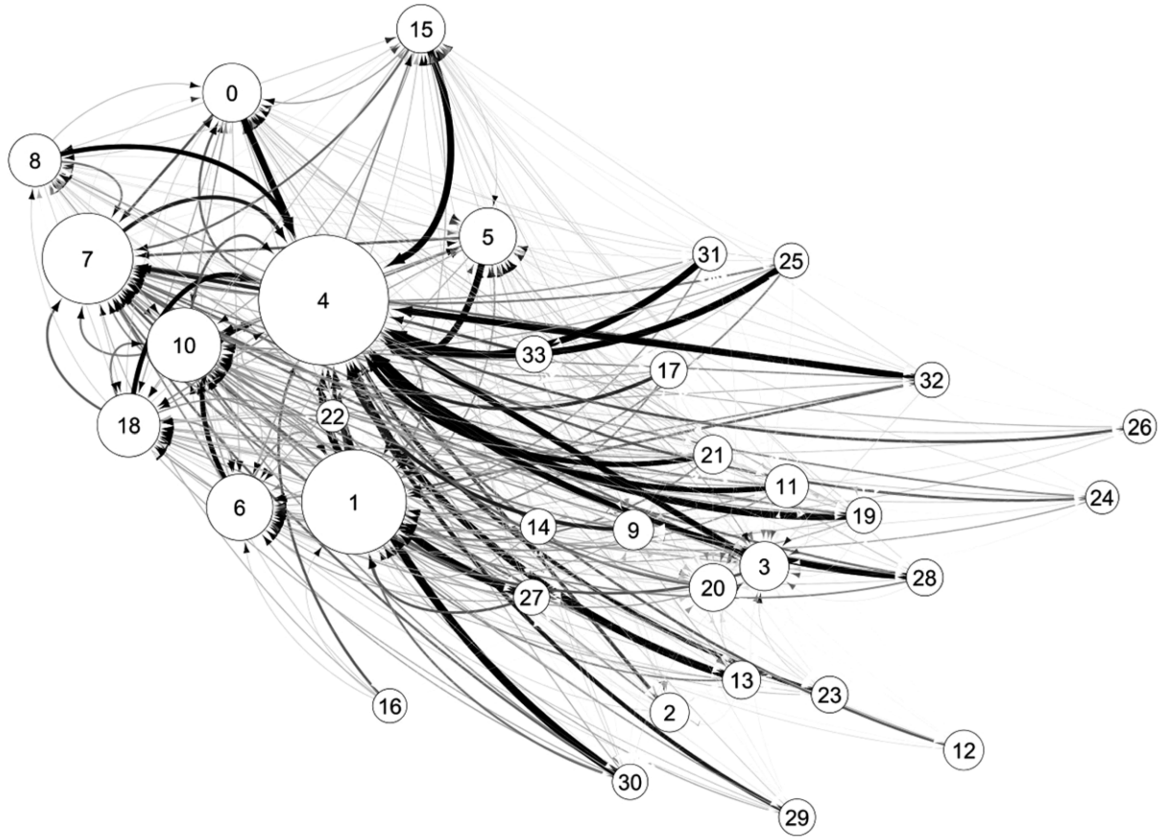


**Fig. 9.** Community structure for the retransmission graph. Nodes represent communities and edges represent the inter-community links. The nodes' size are proportional to the number of people that compound the community and the edge widths are proportional to the number of inter-community links normalized by the size of the community.

users like group 4. On the right side of Fig. 8 we also show how the followers communities retransmit messages from other communities. This behavior has been pointed by Przemyslaw et al. [11] who demonstrated that retweets transcend the friends communities and serve as bridges for messages to spread throughout the network.

On the other hand, we also calculated the community structure for the retransmission graph. In this case we found 34 main communities containing more than 96% of the population. This network showed a completely different mesoscale structure as can be observed in Fig. 9. However, a similar user behavior as the follower network has been detected. Each of these communities contains at its core at least one popular account, like the information producers described above, which is highly followed and retransmitted. In Table 3 we present some information about these popular accounts, like their nicknames and profession. Once again we found communities formed around traditional mass media agents, such as TV Stations, newspapers, journalists and politicians, as well as humorists, civilian activists, student leaders, community managers and micro-bloggers. This result indicates that people behave selectively when retransmitting messages in comparison to just receiving them.

Once again the nodes represent the communities detected and the sizes are proportional to the amount of users compounding each of them. The edges represent the inter-community links, and its width is proportional to the amount of inter-communities links found, normalized by the size of the outgoing community. These links exist due to the fact that some users retransmitted messages from another community. It can be noticed that communities are also asymmetrical when referring to inter-community retransmissions, and also present different profiles. For example, the community number four, which is formed around the Venezuelan micro-blogger @cualrevolucion and Cuban blogger @yoanisanchez, is highly retransmitted by all other communities, while it hardly retransmitted other communities.

## 6. Conclusions

We studied the Venezuelan protest #SOSInternetVE, which took place exclusively on Twitter. We have analyzed the structure and behavior of the participant users based on their information exchange interactions. For this we have constructed two networks to represent the social substratum, where information may flow, and the information diffusion graph, where messages actually travel. Most of the degree distributions in both networks follow power laws and the mean distances between nodes resulted to be very small.

**Table 3**
Most retransmitted account at each retransmission community.

| Community | Popular account | Collective |
|---|---|---|
| 0 | @nelsonbocaranda | Journalist |
| 1 | @rctv_contigo | TV Station |
| 2 | @elnacionalweb | Newspaper |
| 3 | @indiferencia | Community manager |
| 4 | @cualrevolucion, @yoanisanchez | Micro-bloggers |
| 5 | @ucabistas | Student leaders |
| 6 | @erikadlv | Journalist |
| 7 | @vvperiodistas | TV Station |
| 8 | @kikobautista | Journalist |
| 9 | @edoilustrado | Political comics |
| 10 | @globovision | TV Station |
| 11 | @palabrasdecersar | Humorist Micro-blogger |
| 12 | @rmh1947 | Government favorable activist |
| 13 | @leopoldolopez | Politician |
| 14 | @carlossicilia | Humorist |
| 15 | @alberto_ravell | Journalist |
| 16 | @gabycastellanos | Community manager |
| 17 | @ecualink | Ecuadorian magazine |
| 18 | @leonardo_padron | Writer |
| 19 | @EUTrafico | Newspaper |
| 20 | @2010misterchip | Sports journalist |
| 21 | @simplescomillas | Micro-blogger |
| 22 | @ledvarela | Micro-blogger |
| 23 | @oscar_azacon | Journalist |
| 24 | @josbarrios | Micro-blogger |
| 25 | @soymalandro | Humorist Micro-blogger |
| 26 | @mueveteahoracom | Chilean radio station |
| 27 | @jossolavarria | Micro-blogger |
| 28 | @randompiece | Micro-blogger |
| 29 | @unoticias | Newspaper |
| 30 | @manuelrosales | Politician |
| 31 | @blackberryvzla | Smart phone brand |
| 32 | @usemistas | Student leaders |
| 33 | @jorgeramosnews | International Journalist |

Then, based on the network structure, we identified three types of user behavior that determine the dynamics of the information flow: Information Producers, Active Consumers and Passive Consumers. We found some users that cause a lot of activity inside the network, posting a small amount of messages, while others must post lots of messages in order to get retransmitted. We also found a big fraction of very passive users who do not retransmit or get retransmitted at all.

We also carried out a community analysis to describe the mesoscale structure of the networks. We found that people are organized around different collectives. The most central users who conform each of these collectives are very popular and usually they also generate smaller retransmission communities emergent from the propagation dynamic. This shows that people are more selective when it comes to take an active part in the conversation.

Finally, we conclude that although the online social media seems to be a pure social phenomena, traditional media agents still enjoy a lot of power and influence over people, who they use to boost and enhance their messages.

## Acknowledgment

Support from MICINN-Spain under contracts No. MTM2009- 14621, and i-MATH CSD2006-32, is gratefully acknowledged.

## References

[1] S. Wu, J.M. Hofman, W.A. Mason, D.J. Watts, Who says what to whom on Twitter, in: 20th Annual World Wide Web Conference, ACM, Hyderabad, India, 2011.
[2] Javier Borge-Holthoefer, Alejandro Rivero, Iigo Garcia, Elisa Cauhe, Alfredo Ferrer, Dario Ferrer, David Francos, David Iiguez, Maria Pilar Perez, Gonzalo Ruiz, Francisco Sanz, Fermin Serrano, Cristina Vias, Alfonso Tarancon, Yamir Moreno, Structural and dynamical patterns on online social networks: the Spanish May 15th movement as a case study, PLoS One 6 (8) (2011) e23883. Epub.
[3] David Lazer, et al., Computational social science, Science (2009) 721–723.
[4] Akshay Java, Xiaodan Song, Tim Finin, Belle Tseng, Why we Twitter: understanding microblogging usage and communities, in: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, August 12, 2007, San Jose, California, pp. 56–65.
[5] H. Kwak, C. Lee, H. Park, S. Moon, What is Twitter, a social network or a news media? in: Proceedings of WWW, 2010.
[6] B.A. Huberman, D.M. Romero, F. Wu, Social networks that matter: Twitter under the microscope, First Monday 14 (1) (2009).
[7] J. Ratkiewicz, S. Fortunato, A. Flammini, F. Menczer, A. Vespignani, Characterizing and modeling the dynamics of online popularity, Physiscal Review Letters 105 (2010) 158701.

[8] M. Cha, H. Haddadi, F. Benevenuto, K.P. Gummad, Measuring user influence on Twitter: the million follower fallacy, in: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington, DC, 2010, pp. 10–17.
[9] Daniel Gayo-Avello, Nepotistic relationships in Twitter and their impact on rank prestige algorithms, 2010. Arxiv preprint arXiv:1004.0816.
[10] D.M. Romero, W. Galuba, S. Asur, B.A. Huberman, Influence and passivity in social media, in: Proceedings of WWW 2011, ACM, New York, NY, USA, 2011.
[11] Przemyslaw A. Grabowicz, José J. Ramasco, Esteban Moro, Josep M. Pujol, Victor M. Eguiluz, Social features of online networks: the strength of weak ties in online social media, CoRR, Vol. abs/1107.4009, 2011.
[12] Wojciech Galuba, Karl Aberer, Dipanjan Chakraborty, Zoran Despotovic, Wolfgang Kellerer, Outtweeting the Twitterers — predicting information cascades in microblogs, in: Proceedings of the 3rd Conference on Online Social Networks, June 22–25, 2010, USA, p. 3.
[13] R. Albert, A.-L. Barabasi, Statistical mechanics of complex networks, Reviews of Modern Physics 74 (2002) 47.
[14] comScore, Inc. Social networking on-the-go: U.S. mobile social media audience grows 37 percent in the past year, Press Release, 2011, http://www.comscore.com/.
[15] Jaewon Yang, Jure Leskovec, Patterns of temporal variation in online media, in: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, February 09–12, 2011, Hong Kong, China.
[16] Sitaram Asur, Bernardo A. Huberman, Gabor Szabo, Chunyan Wang, Trends in social media: persistence and decay, 2011. Available at SSRN: http://ssrn.com/abstract=1755748.
[17] M.E.J. Newman, Power laws, Pareto distributions and Zipf's law, Contemporary Physics 46 (5) (2006) 323–351.
[18] D.J. Watts, S.H. Strogatz, Collective dynamics of small world networks, Nature 393 (1998) 440–442.
[19] S. Milgram, The small world problem, Psychology Today 2 (1967) 60–67.
[20] B. Gonalves, N. Perra, A. Vespignani, Modeling users' activity on Twitter networks: validation of Dunbar's number, PLoS One 6 (8) (2011) e22656.
[21] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, Proceedings of the National Academy of Sciences 99 (12) (2002) 7821–7826.
[22] V.D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, Journal of Statistical Mechanics (2008) P10008.